

In the Claims

1. (Currently amended) One or more computer-readable media having stored thereon a plurality of instructions that, when executed by one or more processors of a computer, cause the one or more processors to:

store data indicative of a data transmission rate for each of previous server operations;

generate, by averaging the stored data transmission rates, a bandwidth value indicating an average bandwidth used by a server to perform the previous server operations;

receive a request for the server to perform an additional server operation;

compare the bandwidth value to a first threshold and a second threshold;

and

restrict the request in a first manner if the bandwidth value exceeds the first threshold but does not exceed the second threshold, and restrict the request in a second manner if the bandwidth value exceeds the second threshold;[[.]]

wherein to restrict the request in the first manner is to delay the request;

wherein to restrict the request in the second manner is to block the request.

2. (Canceled).

3. (Currently amended) One or more computer-readable media having stored thereon a plurality of instructions that, when executed by one or more processors of a computer, cause the one or more processors to:

generate a bandwidth value indicating an average bandwidth used by a server to perform previous server operations;

receive a request for the server to perform an additional server operation;

compare the bandwidth value to a first threshold and a second threshold;

and

restrict the request in a first manner if both the bandwidth value exceeds the first threshold and the additional server operation is of a first type, and restrict the request in a second manner if the bandwidth value exceeds the second threshold;[[.]]

wherein to restrict the request in the first manner is to delay the request;

wherein to restrict the request in the second manner is to block the request.

4. (Original) One or more computer-readable media as recited in claim 3, wherein the first type comprises a read operation.

5. (Original) One or more computer-readable media as recited in claim 3, wherein the first type comprises a write operation.

6. (Original) One or more computer-readable media as recited in claim 3, wherein the first type comprises a transmit operation.

1 7. (Original) One or more computer-readable media as recited in claim
2 1, wherein the previous server operations and the additional server operation
3 include one or more of read operations, write operations, and transmit operations.

4
5 8. (Previously presented) One or more computer-readable media
6 having stored thereon a plurality of instructions that, when executed by one or
7 more processors of a computer, cause the one or more processors to:

8 generate a bandwidth value indicating an average bandwidth used by a
9 server to perform previous server operations;

10 receive a request for the server to perform an additional server operation;

11 compare the bandwidth value to a threshold;

12 restrict the request in a first manner if the bandwidth value exceeds the
13 threshold, wherein to restrict the request in the first manner is to delay the request;

14 compare the bandwidth value to another threshold, wherein the threshold is
15 less than the another threshold; and

16 restrict the request in a second manner if the bandwidth value exceeds the
17 another threshold, wherein to restrict the request in the second manner is to block
18 the request.

19
20 9. (Canceled).

21
22 10. (Canceled).

23
24 11. (Canceled).

25

1 12. (Canceled).

2
3 13. (Canceled).

4
5 14. (Previously presented) One or more computer-readable media
6 having stored thereon a plurality of instructions that, when executed by one or
7 more processors of a computer, cause the one or more processors to:
8 generate a bandwidth value indicating an average bandwidth used by a
9 server to perform previous server operations by:

10 generating a value for each of the previous server operations by,

11 identifying a time interval duration between a start time of the
12 previous server operation and an end time of the previous server
13 operation,

14 identifying a number of bytes transferred for the previous
15 server operation, and

16 dividing the number of bytes by the time interval duration;

17 and

18 dividing a sum of the values of the previous server operations by the
19 number of previous server operations;

20 receive a request for the server to perform an additional server operation;

21 compare the bandwidth value to a threshold; and

22 restrict the request in a first manner if the bandwidth value exceeds the
23 threshold.

1 15. (Currently amended) A method comprising:
2 receiving a request to perform a server operation;
3 storing data indicative of a data transmission rate for each of previous
4 server operations;
5 generating an average bandwidth by averaging the stored data transmission
6 rates; and
7 restricting performance of the request based at least in part on the average
8 bandwidth, wherein the restricting comprises:
9 comparing the average bandwidth to at least one of a first threshold
10 and a second threshold;
11 restricting the request in a first manner if the average bandwidth
12 exceeds the first threshold but does not exceed the second threshold; and
13 restricting the request in a second manner if the average bandwidth
14 exceeds the second threshold;[[.]]
15 wherein ~~to restricting~~ the request in the first manner comprises ~~is to~~
16 delaying the request;
17 wherein ~~to restricting~~ the request in the second manner comprises ~~is~~
18 to blocking the request.

19
20 16. (Canceled).
21
22
23
24
25

1 17. (Previously presented) A method as recited in claim 15, wherein the
2 restricting further comprises restricting the request in the first manner if both the
3 average bandwidth exceeds the first threshold and the server operation is of a first
4 type.

5
6 18. (Original) A method as recited in claim 17, wherein the first type
7 comprises a read operation.

8
9 19. (Original) A method as recited in claim 17, wherein the first type
10 comprises a write operation.

11
12 20. (Original) A method as recited in claim 17, wherein the first type
13 comprises a transmit operation.

14
15 21. (Canceled).

16
17 22. (Canceled).

18
19 23. (Canceled).

20
21 24. (Currently amended) A host system comprising:
22 at least one network server;
23 an asynchronous thread queue to receive a request, from a client process, to
24 be performed by one of the at least one network server; and
25

1 a bandwidth throttling system, coupled to the asynchronous thread queue,
2 to determine whether performance of the request by the one of the at least one
3 network server is to be restricted in a first manner or a second manner based at
4 least in part on both an average bandwidth used by the one of the at least one
5 network server in performing previous server operations and on a type of the
6 request, wherein the bandwidth throttling system is further to:

7 store data indicative of a data transmission rate for each of the
8 previous server operations; and

9 generate the average bandwidth by averaging the stored data
10 transmission rates; and

11 wherein to restrict the request in the first manner is to delay the
12 request;

13 wherein to restrict the request in the second manner is to block the
14 request;
15

16 25. (Canceled).

17
18 26. (Original) A host system as recited in claim 24, further comprising
19 an ancillary function driver, coupled to the asynchronous thread queue, to couple
20 the host system to a network.
21

22 27. (Previously presented) A host system comprising:
23 at least one network server;
24 an asynchronous thread queue to receive a request, from a client process, to
25 be performed by one of the at least one network server; and

1 a bandwidth throttling system, coupled to the asynchronous thread queue,
2 to determine whether performance of the request by the one of the at least one
3 network server is to be restricted based at least in part on an average bandwidth
4 used by the one of the at least one network server in performing previous
5 operations;

6 wherein the at least one network server comprises a plurality of network
7 servers, and wherein the bandwidth throttling system determines whether
8 performance of a request by a particular one of the plurality of network servers is
9 to be restricted based on the average bandwidth used by that particular network
10 server in performing previous operations and independent of the average
11 bandwidth used by other network servers of the plurality of network servers in
12 performing previous operations.

13
14 28. (Previously presented) A host system comprising:

15 at least one network server;

16 an asynchronous thread queue to receive a request, from a client process, to
17 be performed by one of the at least one network server; and

18 a bandwidth throttling system, coupled to the asynchronous thread queue,
19 to determine whether performance of the request by the one of the at least one
20 network server is to be restricted based at least in part on an average bandwidth
21 used by the one of the at least one network server in performing previous
22 operations;

23

24

25

wherein the bandwidth throttling system comprises:

a measurement subsystem to compute the average bandwidth used by the one of the at least one network server in performing previous operations; and

a control subsystem, coupled to the measurement subsystem, to make the determination and communicate how the asynchronous thread queue is to restrict performance of the request.

29. (Currently amended) A method comprising:

storing data indicative of a data transmission rate for each of previous server operations;

generating, by averaging the stored data transmission rates, a value indicating a bandwidth used by a server to perform the previous server operations;

receiving a request for the server to perform an additional server operation; and

restricting the request in a first manner if the additional server operation is of a first type and the value exceeds a first threshold but does not exceed a second threshold, and restricting the request in a second manner if the value exceeds the second threshold;

wherein to restrict the request in the first manner is to delay the request;

wherein to restrict the request in the second manner is to block the request;

~~wherein to restrict the request in the first manner is to delay the request~~

~~wherein to restrict the request in the second manner is to block the request~~

1 30. (Canceled).

2
3 31. (Previously presented) A method comprising:

4 generating a value indicating a bandwidth used by a server to perform
5 previous server operations;

6 receiving a request for the server to perform an additional server operation;

7 restricting the request in a first manner if the value exceeds a first threshold
8 but does not exceed a second threshold, and restricting the request in a second
9 manner if the value exceeds the second threshold;

10 generating another value indicating bandwidth used by another server to
11 perform other previous server operations;

12 receiving another request for the another server to perform another
13 additional server operation; and

14 restricting the other request in the first manner if the value exceeds a third
15 threshold but does not exceed a fourth threshold, and restricting the other request
16 in the second manner if the value exceeds the fourth threshold.

17
18 32. (Previously presented) A method as recited in claim 31, wherein the
19 first threshold is different than the third threshold, and wherein the second
20 threshold is different than the fourth threshold.

21
22 33. (Canceled).

23
24 34. (Canceled).

25

1 35. (Canceled).

2
3 36. (Canceled).

4
5 37. (Original) A method as recited in claim 29, wherein the value
6 indicates an average bandwidth used by the server.

7
8 38. (Original) A method as recited in claim 29, wherein the restricting
9 comprises restricting the request in the first manner or the second manner only if
10 the request is of a first type.

11
12 39. (Original) A method as recited in claim 38, wherein the first type
13 comprises a read operation.

14
15 40. (Original) A method as recited in claim 38, wherein the first type
16 comprises a write operation.

17
18 41. (Original) A method as recited in claim 38, wherein the first type
19 comprises a transmit operation.

20
21 42. (Original) One or more computer-readable memories comprising
22 computer-readable instructions that, when executed by a processor, direct a
23 computer system to perform the method as recited in claim 29.
24
25

1 43. (Currently amended) One or more computer-readable media having
2 stored thereon a plurality of instructions that, when executed by one or more
3 processors of a computer, cause the one or more processors to:

4 store data indicative of a data transmission rate for each of previous server
5 operations;

6 generate, by averaging the stored data transmission rates, a value indicating
7 a bandwidth used by a server to perform the previous server operations;

8 receive a request for the server to perform an additional server operation;

9 compare the value to at least one of a first threshold and a second threshold;

10 restricting performance of a request based at least in part on an average
11 bandwidth used in performing said previous server operations; and

12 delay the request if the value exceeds the first threshold but does not exceed
13 the second threshold, and block the request if the value exceeds the second
14 threshold.

15
16 44. (Canceled).

17
18 45. (Original) One or more computer-readable media as recited in claim
19 43, wherein the value indicates an average bandwidth used by the server.

1 46. (Currently amended) One or more computer-readable media having
2 stored thereon a plurality of instructions that, when executed by one or more
3 processors of a computer, cause the one or more processors to:

4 generate a value indicating a bandwidth used by a server to perform
5 previous server operations by averaging stored data transmission rates indicative
6 of data transmission rates of previous server operations;

7 receive a request for the server to perform an additional server operation;

8 compare the value to at least one of a first threshold and a second threshold;

9 and

10 restricting performance of a request based at least in part on an average
11 bandwidth used in performing said previous server operations; and

12 delay the request only if the request is of a first type and if the value
13 exceeds the first threshold but does not exceed the second threshold, and block the
14 request only if the request is of the first type and if the value exceeds the second
15 threshold.

16
17 47. (Original) One or more computer-readable media as recited in claim
18 46, wherein the first type comprises a read operation.

19
20 48. (Original) One or more computer-readable media as recited in claim
21 46, wherein the first type comprises a write operation.

22
23 49. (Original) One or more computer-readable media as recited in claim
24 46, wherein the first type comprises a transmit operation.

25

1 50. (Canceled).

2
3 51. (Previously presented) A host system comprising:

4 at least one network server;

5 an asynchronous thread queue to receive a request, from a client process, to
6 be performed by one of the at least one network server; and

7 a bandwidth throttling system, coupled to the asynchronous thread queue,
8 to compare a value indicating a bandwidth used by the one of the at least one
9 network server to perform previous server operations to at least one of a first
10 threshold and a second threshold, to restrict the request in a first manner if the
11 value exceeds the first threshold but does not exceed the second threshold, and to
12 restrict the request in a second manner if the value exceeds the second threshold;

13 wherein the bandwidth throttling system is further to:

14 store data indicative of a data transmission rate for each of the
15 previous server operations; and

16 generate the value by averaging the stored data transmission rates.

17
18 52. (Previously presented) A host system comprising:

19 at least one network server;

20 an asynchronous thread queue to receive a request, from a client process, to
21 be performed by one of the at least one network server; and

22 a bandwidth throttling system, coupled to the asynchronous thread queue,
23 to compare a value indicating a bandwidth used by the one of the at least one
24 network server to perform previous server operations to at least one of a first
25 threshold and a second threshold, to restrict the request in a first manner if the

1 value exceeds the first threshold but does not exceed the second threshold, and to
2 restrict the request in a second manner if the value exceeds the second threshold;

3 further comprising a plurality of network servers, and wherein the
4 bandwidth throttling system determines whether to restrict a request for a
5 particular one of the plurality of network servers based on the value indicating
6 bandwidth used by that particular network server in performing the previous
7 server operations and independent of bandwidth used by other network servers of
8 the plurality of network servers in performing the previous server operations.

9
10 53. (Previously presented) A host system comprising:

11 at least one network server;

12 an asynchronous thread queue to receive a request, from a client process, to
13 be performed by one of the at least one network server; and

14 a bandwidth throttling system, coupled to the asynchronous thread queue,
15 to compare a value indicating a bandwidth used by the one of the at least one
16 network server to perform previous server operations to at least one of a first
17 threshold and a second threshold, to restrict the request in a first manner if the
18 value exceeds the first threshold but does not exceed the second threshold, and to
19 restrict the request in a second manner if the value exceeds the second threshold;

20 wherein the bandwidth throttling system comprises:

21 a measurement subsystem to compute, as the value, an average
22 bandwidth used by the network server in performing the previous server
23 operations; and
24
25

1 a control subsystem, coupled to the measurement subsystem, to
2 make the determination and communicate how the asynchronous thread
3 queue is to restrict performance of the request.

4
5 54. (Original) One or more computer-readable media having stored
6 thereon a plurality of instructions that, when executed by one or more processors
7 of a computer, cause the one or more processors to:

8 determine a presently used bandwidth for each of at least one network
9 server by way of a data transmission rate measurement during execution of an
10 operation for each of said at least one network server that includes:

11 storing data indicative of said data transmission rate measurement of
12 said operation for n last most recently executed operations for each of said
13 at least one network server, wherein n is a positive integer greater than 2,
14 and

15 generating data indicative of an effective presently used bandwidth
16 for each of said at least one network server wherein said effective presently
17 used bandwidth includes an average of said data transmission rate
18 measurement over said n last most recently executed operations;

19 effect provision of a plurality of classes of service provided by one of said
20 at least one network server in a first manner, in response to said effective presently
21 used bandwidth for said one of said at least one network server that exceeds a first
22 threshold; and

23 effect provision of said plurality of classes of service provided by said one
24 of said at least one network server in a second manner that differs from said first
25 manner, in response to said effective presently used bandwidth for said one of said

1 at least one network server that exceeds a second threshold that is greater than said
2 first threshold.

3
4 55. (Previously presented) One or more computer-readable media
5 having stored thereon a plurality of instructions that, when executed by one or
6 more processors of a computer, cause the one or more processors to:

7 store data indicative of a predetermined allocated data transmission
8 bandwidth for each of a plurality of network servers, each of said plurality of
9 network servers providing a plurality of classes of service;

10 determine an effective bandwidth for each of said plurality of network
11 servers, comprising:

12 calculating a bandwidth for each operation performed by each of
13 said plurality of network servers,

14 tabulating a count of bandwidth used by each of said plurality of
15 network servers in each of a plurality of last time intervals, and

16 averaging said tabulated count of bandwidth to obtain a value
17 indicative of said effective bandwidth;

18 delay, in response to said effective bandwidth for a one of said network
19 servers substantially corresponding to said predetermined allocated data
20 transmission bandwidth for said one network server, for delaying a first subset of
21 said plurality of classes of service provided by said one network server; and

22 reject, in response to said effective bandwidth for said one network server
23 exceeding said predetermined allocated data transmission bandwidth for said one
24 network server, requests for said first subset of classes of service provided by said
25

1 one network server and for delaying service for a second subset of said plurality of
2 classes of service provided by said one network server.

3
4 56. (Original) One or more computer-readable media having stored
5 thereon a plurality of instructions that, when executed by one or more processors
6 of a computer, cause the one or more processors to:

7 determine a presently used bandwidth for each of a plurality of network
8 servers by way of a data transmission rate measurement taken during execution of
9 an operation for each of said plurality of network servers, said determining
10 including:

11 storing data indicative of a predetermined allocated data
12 transmission bandwidth for each of said plurality of network servers, and

13 storing data indicative of a first threshold, wherein said first
14 threshold includes an indication of a differential from said predetermined
15 allocated data transmission bandwidth for each of said plurality of network
16 servers; and

17 effect provision of a plurality of classes of service provided by a first one of
18 said plurality of network servers in a manner that is individually defined for a
19 plurality of successively greater thresholds in response to said effective presently
20 used bandwidth of said first one of said plurality of network servers exceeding one
21 of said plurality of successively greater thresholds, wherein said manner defined
22 for said first one of said plurality of network servers differs from every other
23 manner defined for other ones of said plurality of network servers.